

ABSTRACT

A system and method for data element extraction from hierarchically structured documents. According to one embodiment, a generator extracts a plurality of data elements from a stored hierarchical data file that is in a first format, in accordance with a data set that is unique to the particular hierarchical data file, and arranges the plurality of data elements in a second format so that the arrangement provides the plurality of data elements with a document context, and so that the data elements can be further arranged in a third format without using the data set.